

# SUBSTITUTE SEQUENCE LISTING

<110> DE WAARD, Michel  
 DUPUIS, Alain  
 GRUNWALD, Didier  
 SANDOZ, Guillaume

<120> CHIMERIC PROTEIN FOR THE SCREENING OF AGONISTS AND ANTAGONISTS OF  
 CELL SIGNALLING PATHWAYS THAT ARE DEPENDENT ON G-PROTEIN-COUPLED  
 RECEPTORS

<130> 273623US0XPCT

<140> 10/540,247  
 <141> 2005-06-22

<150> PCT/FR03/003860  
 <151> 2003-12-22

<150> FR 02/16576  
 <151> 2002-12-23

<160> 14

<170> PatentIn version 3.3

<210> 1  
 <211> 58  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA

<400> 1  
 tttggtacca tggatgacga ctctacgtg cccgggtttg aggactcgga ggcggggt 58

<210> 2  
 <211> 48  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA

<400> 2  
 gcggaattcg tagctgtcct taggccaagg ccggttacgc tgccagtt 48

<210> 3  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA

<400> 3  
 ggggaattcg ccaaagaaag ggagcgggtg gagaac 36

<210> 4  
 <211> 51  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA

<400> 4  
 tttgaattct tactgagttt tgaccatgcg acggatgtag aaacgcattc t 51

<210> 5  
 <211> 7242  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA

<400> 5  
 atggatgacg actcctacgt gcccgggttt gaggactcgg aggcgggttc agccgactcc 60  
 tacaccagcc gccctctctt ggactcagac gtttccctgg aggaggaccg ggagagtgcc 120  
 cggcgagaag tggagagtca ggctcagcag cagctggaaa gagccaagca caaacctgtg 180  
 gcatttgctg tgaggaccaa tgtcagctac tgtggagtgc tggatgagga atgccagtc 240  
 cagggctctg gagtcaactt cgaggccaaa gattttctgc acattaaaga gaagtacagc 300  
 aatgactggg ggatcgggag gctagtgaag gaagggtggc atattgcctt catccccagc 360  
 ccccaacgcc tggagagcat ccggctcaaa caggaacaga aggccaggag atccgggaac 420  
 ccttccagcc tgagtacat tggcaaccga cgttcccctc ctccatctct agccaagcag 480  
 aagcaaaagc aggcggaaca tgtccccccg tatgatgtgg tgccctccat gcggcctgtg 540  
 gtgctggtgg gaccctctct gaaagggtat gaggtcacag acatgatgca gaaggctctc 600  
 ttcgacttcc ttaaacacag gtttgatggc aggatctcca tcaccgcgt cacggctgac 660  
 ctctcactgg ccaagcgctc tgtgctcaac aatcctggca agaggaccat catcgagcgc 720  
 tcttctgccc gctccagcat tgctgaggtg cagagtgaga ttgagcgcat attcgagctg 780  
 gccaaatccc tgcagctagt ggtgttggat gctgacacca tcaaccaccc agcacagcta 840  
 gccaaagacct cactggcccc catcatcgtc ttcgtcaaag tgtcctcgcc aaaggctactg 900  
 cagcgactga tccgctccag ggggaagtcc cagatgaagc acctcactgt acagatgatg 960

gcatatgata agctgggttca gtgcccacct gagtcatttg atgtgattct ggatgagaac	1020
cagctggacg acgcctgtga gcacctagct gaatacctag aggtttactg gcgcgctacc	1080
caccacccag caccggggccc cgggatgctg ggtccgcccga gtgccatccc tggacttcag	1140
aaccagcagc tgctggggga gcgaggtgag gagcattcac ccctggagcg ggacagtttg	1200
atgccctcgg atgaggccag tgagagctcc cgccaggcct ggaccggatc ttcacagcgc	1260
agctcccgcc atctggagga ggactatgca gatgcctacc aggacctgta ccagcctcac	1320
cgtcaacaca cctcggggct acccagtgct aacgggcatg accccaaga cgggtccta	1380
gccaggact cggagcatga ccacaatgac cggaactggc agcgtaaccg gccttggcct	1440
aaggacagct acgaattcgc caaagaaagg gagcgggtgg agaaccggcg cgcattcctg	1500
aagctgcggc ggcagcagca gattgaacgc gagctcaacg ggtacatgga gtggatctca	1560
aaagcagaag aggtgatcct cgcagaggac gagaccgacg tggagcagag acatcccttt	1620
gatggagctc tgcgagagac cactatcaag aagagcaaga cggacctgct ccaccagag	1680
gaggcggagg atcagctggc cgacatcgcc tccgtggggc ctccttttgc ccgagccagc	1740
attaaaagtg ccaagctgga gaactcgagt tttttccaca aaaaagagag gagaatgcgt	1800
ttctacatcc gtcgcatggc caaaactcag taagaattct gcagatatcc atcacactgg	1860
cggccgctcg agcatgcac tagaggggcc tattctatag tgtcacctaa atgctagagc	1920
tcgctgatca gcctcgactg tgctttctag ttgccagcca tctgttggtt gccctcccc	1980
cgtgccttcc ttgacctgg aaggtgccac tcccactgtc ctttcctaataaaaatgagga	2040
aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga	2100
cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat	2160
ggcttctgag gcggaaagaa ccagctgggg ctctaggggg tatccccacg cgccctgtag	2220
cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag	2280
cgccctagcg cccgctcctt tcgctttctt ccttccctt ctcgccacgt tcgccggctt	2340
tccccgtcaa gctctaaatc ggggcatccc tttagggttc cgatttagtg ctttacggca	2400
cctcgacccc aaaaaacttg attaggggtga tggttcacgt agtgggcat cgccctgata	2460
gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggac tcttgttcca	2520
aactggaaca aactcaacc ctatctcggt ctattctttt gatattataag ggattttggg	2580
gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattaatt	2640

ctgtggaatg tgtgtcagtt aggggtgtgga aagtccccag gctccccagg caggcagaag	2700
tatgcaaagc atgcatctca attagtcagc aaccagggtgt ggaaagtccc caggctcccc	2760
agcaggcaga agtatgcaaa gcatgcatct caattagtcg gcaaccatag tccccgccct	2820
aactccgccc atcccccccc taactccgcc cagttccgcc cattctccgc cccatggctg	2880
actaattttt tttatttatg cagaggccga ggccgcctct gcctctgagc tattccagaa	2940
gtagtgagga ggcttttttg gaggcctagg cttttgcaaa aagctcccg gagcttgtat	3000
atccattttc ggatctgac aagagacagg atgaggatcg ttctgcatga ttgaacaaga	3060
tggattgcac gcaggttctc cgcccgcttg ggtggagagg ctattcggct atgactgggc	3120
acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcccc	3180
ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg acgaggcagc	3240
gcggctatcg tggctggcca cgacgggctg tccttgcgca gctgtgctcg acgttgtcac	3300
tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc tcctgtcatc	3360
tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc ggctgcatac	3420
gcttgatccg gctacctgcc cattcgacca ccaagcgaac catcgcatcg agcgagcacg	3480
tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc atcaggggct	3540
cgcgccagcc gaactgttcg ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt	3600
cgtgacccat ggcatgctt gcttgccgaa tatcatgggtg gaaaatggcc gcttttctgg	3660
attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggctac	3720
ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg tgctttacgg	3780
tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg agttcttctg	3840
agcgggactc tggggttcga aatgaccgac caagcgacgc ccaacctgcc atcacgagat	3900
ttcgattcca ccgccgcctt ctatgaaagg ttgggcttcg gaatcgtttt ccgggacgcc	3960
ggctggatga tcctccagcg cggggatctc atgctggagt tcttcgcccc cccaacttg	4020
tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt cacaataaaa	4080
gcattttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt atcttatcat	4140
gtctgtatac cgtcgacctc tagctagagc ttggcgtaat catggtcata gctgtttcct	4200
gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt	4260
aaagcctggg gtgcctaata agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc	4320
gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg	4380

agaggcggtt	tgcgtattgg	gcgctcttcc	gcttcctcgc	tactgactc	gctgcgctcg	4440
gtcgttcggc	tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatcag	gttatccaca	4500
gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaag	gccagcaaaa	ggccaggaac	4560
cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	4620
aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	gactataaag	ataccaggcg	4680
tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	4740
ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	aatgctcacg	ctgtaggtat	4800
ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	4860
cccgaccgct	gcgccttata	cggtaactat	cgtcttgagt	ccaaccgggt	aagacacgac	4920
ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	4980
gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	agtatttggt	5040
atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	5100
aaacaaacca	ccgctggtag	cggtgggtttt	tttgtttgca	agcagcagat	tacgcgcaga	5160
aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaac	5220
gaaaactcac	gttaagggat	tttggtcatg	agattatcaa	aaaggatctt	cacctagatc	5280
cttttaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttggtct	5340
gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	5400
tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	5460
ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	tttatcagca	5520
ataaaccagc	cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	5580
atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	5640
cgcaacgttg	ttgccattgc	tacaggcatc	gtggtgtcac	gctcgtcggt	tggtatggct	5700
tcattcagct	ccggttccca	acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	5760
aaagcggtta	gctccttcgg	tcctccgatc	gttgtcagaa	gtaagttggc	cgcagtgtta	5820
tactcatgg	ttatggcagc	actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	5880
ttttctgtga	ctgggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	5940
agttgctctt	gcccggcgctc	aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	6000
gtgctcatca	ttggaaaacg	ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	6060

```

agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 6120
accagcgttt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa ggaataagg 6180
gcgacacgga aatgttgaat actcatactc ttcttttttc aatattattg aagcatttat 6240
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6300
gggggtccgc gcacatttcc ccgaaaagtg ccacctgacg tcgacggatc gggagatctc 6360
ccgatcccct atgggtcgact ctcaagtaca tctgctctga tgccgcatag ttaagccagt 6420
atctgctccc tgcttgtgtg ttggaggctg ctgagtagtg cgcgagcaaa atttaagcta 6480
caacaaggca aggcttgacc gacaattgca tgaagaatct gcttaggggt aggcggtttg 6540
cgctgcttcg cgatgtacgg gccagatata cgcgttgaca ttgattattg actagttatt 6600
aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc cgcgttacat 6660
aacttacggt aaatggcccg cctggctgac cgccaacga ccccgccca ttgacgtcaa 6720
taatgacgta tgttcccata gtaacgccaa tagggacttt ccattgacgt caatgggtgg 6780
actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg ccaagtacgc 6840
cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag tacatgacct 6900
tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt accatggtga 6960
tgcggttttg gcagtacatc aatgggcgtg gatagcgggt tgactcacgg ggatttccaa 7020
gtctccacc cattgacgtc aatgggagtt tgttttggca ccaaatcaa cgggactttc 7080
caaatgtcg taacaactcc gcccattga cgcaaatggg cggtaggcgt gtacggtggg 7140
aggtctatat aagcagagct ctctggctaa ctagagaacc cactgcttac tggcttatcg 7200
aaattaatac gactcactat agggagaccc aagcttggtg cc 7242

```

```

<210> 6
<211> 39
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic DNA

```

```

<400> 6
agccgtacgc gatcgcatct ctagccaagc agaagcaaa 39

```

```

<210> 7
<211> 42
<212> DNA
<213> Artificial Sequence

```

<220>  
 <223> Synthetic DNA

<400> 7  
 cccgttaacc ccactagtct gagttttgac catgcgacgg at 42

<210> 8  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA

<400> 8  
 gggactagta tggtagagcaa gggcgaggag ctg 33

<210> 9  
 <211> 34  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA

<400> 9  
 cccgttaact gccgagagtg atcccggcgg cggt 34

<210> 10  
 <211> 7242  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA

<400> 10  
 atggatgacg actcctacgt gcccggtttt gaggactcgg aggcgggttc agccgactcc 60  
 tacaccagcc gcccctctct ggactcagac gtttccctgg aggaggaccg ggagagtgcc 120  
 cggcgagaag tggagagtca ggctcagcag cagctggaaa gagccaagca caaacctgtg 180  
 gcatttgctg tgaggaccaa tgtcagctac tgtggagtgc tggatgagga atgcccagtc 240  
 cagggtctctg gagtcaactt cgaggccaaa gattttctgc acattaaaga gaagtacagc 300  
 aatgactggg ggatcgggag gctagtgaaa gaagggtggcg atattgcctt catccccagc 360  
 ccccaacgcc tggagagcat ccggctcaaa caggaacaga aggccaggag atccgggaac 420  
 ccttccagcc tgagtgacat tggcaaccga cgttcccctc ctccatctct agccaagcag 480

aagcaaaagc aggcggaaca tgtcccccg tatgatgtgg tgccctccat gcggcctgtg	540
gtgctggtgg gaccctctct gaaaggttat gaggtcacag acatgatgca gaaggctctc	600
ttcgacttcc ttaaacacag gtttgatggc aggatctcca tcaccgcgt cacggctgac	660
ctctcactgg ccaagcgctc tgtgctcaac aatcctggca agaggaccat catcgagcgc	720
tcttctgccc gctccagcat tgctgaggtg cagagtgaga ttgagcgcat attcgagctg	780
gccaaatccc tgcagctagt ggtggtggat gctgacacca tcaaccaccc agcacagcta	840
gccaaagacct cactggcccc catcatcgtc ttcgtcaaag tgcctcggc aaagggtactg	900
cagcgactga tccgctccag ggggaagtcc cagatgaagc acctcactgt acagatgatg	960
gcatatgata agctggttca gtgcccacct gagtcatttg atgtgattct ggatgagaac	1020
cagctggacg acgctgtga gcacctagct gaatacctag aggtttactg gcgcgtacc	1080
caccacccag caccgggccc cgggatgctg ggtccgcca gtgccatccc tggacttcag	1140
aaccagcagc tgctggggga gcgaggtgag gagcattcac ccctggagcg ggacagtttg	1200
atgccctcgg atgaggccag tgagagctcc cgccaggcct ggaccggatc ttcacagcgc	1260
agtcccgcc atctggagga ggactatgca gatgcctacc aggacctgta ccagcctcac	1320
cgtcaacaca cctcggggct acccagtgtc aacgggcatg accccaaga ccggctccta	1380
gccaggact cggagcatga ccacaatgac cggaactggc agcgtaaccg gccttggcct	1440
aaggacagct acgaattcgc caaagaaagg gagcgggtgg agaaccggcg cgcattcctg	1500
aagctgcggc ggcagcagca gattgaacgc gagctcaacg ggtacatgga gtggatctca	1560
aaagcagaag aggtgatcct cgcagaggac gagaccgacg tggagcagag acatcccttt	1620
gatggagctc tgcggagagc cactatcaag aagagcaaga cggacctgct ccaccagag	1680
gaggcggagg atcagctggc cgacatcgcc tccgtggggg ctccctttgc ccgagccagc	1740
attaaaagtg ccaagctgga gaactcgagt tttttccaca aaaaagagag gagaatgcgt	1800
ttctacatcc gtcgcatggc caaaactcag taagaattct gcagatatcc atcacactgg	1860
cggccgctcg agcatgcac tagaggggcc tattctatag tgtcacctaa atgctagagc	1920
tcgctgatca gcctcgactg tgccttctag ttgccagcca tctgttgttt gccctcccc	1980
cgtgccttcc ttgacctgg aagtgccac tcccactgtc ctttcctaat aaaatgagga	2040
aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga	2100
cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat	2160
ggcttctgag gcggaaagaa ccagctgggg ctctaggggg tatccccacg cgccctgtag	2220



cggcgcatta agcgcggcgg gtgtgggtgg	tacgcgcagc gtgaccgcta cacttgccag	2280
cgccctagcg cccgctcctt tcgctttcct	cccttccttt ctgcgccacgt tcgcgggctt	2340
tccccgtcaa gctctaaatc ggggcatccc	tttaggggtc cgatttagtg ctttacggca	2400
cctcgacccc aaaaaacttg attaggggtga	tggttcacgt agtgggcat cgccctgata	2460
gacgggtttt cgccctttga cgttggagtc	cacgttcctt aatagtggac tcttgttcca	2520
aactggaaca aactcaacc ctatctcggt	ctattccttt gatttataag ggattttggg	2580
gatttcggcc tattgggttaa aaaatgagct	gatttaacaa aaatttaacg cgaattaatt	2640
ctgtggaatg tgtgtcagtt aggggtgtgga	aagtccccag gctccccagg caggcagaag	2700
tatgcaaagc atgcatctca attagtcagc	aaccagggtg ggaaagtccc caggctcccc	2760
agcaggcaga agtatgcaaa gcatgcatct	caattagtca gcaaccatag tcccgcacct	2820
aactccgccc atcccgcctt taactccgcc	cagttccgcc cattctccgc cccatggctg	2880
actaattttt tttatttatg cagaggccga	ggccgcctct gcctctgagc tattccagaa	2940
gtagtgagga ggcttttttg gaggcctagg	cttttgcaaa aagctcccgg gagcttgtat	3000
atccattttc ggatctgac aagagacagg	atgaggatcg tttcgcatga ttgaacaaga	3060
tggattgcac gcaggttctc cgcccgcttg	ggtggagagg ctattcggct atgactgggc	3120
acaacagaca atcggtgct ctgatgccgc	cgtgttcagg ctgtcagcgc aggggcgccc	3180
ggttcttttt gtcaagaccg acctgtccgg	tgccctgaat gaactgcagg acgaggcagc	3240
gcggctatcg tggctggcca cgacgggcgt	tccttgcgca gctgtgctcg acgttgtcac	3300
tgaagcggga agggactggc tgctattggg	cgaagtgccg gggcaggatc tcctgtcatc	3360
tcaccttgct cctgccgaga aagtatccat	catggctgat gcaatgcggc ggctgcatac	3420
gcttgatccg gctacctgcc cattcgacca	ccaagcgaaa catcgcatcg agcgagcacg	3480
tactcggatg gaagccggtc ttgtcgatca	ggatgatctg gacgaagagc atcaggggct	3540
cgcgccagcc gaactgttcg ccaggctcaa	ggcgcgcata cccgacggcg aggatctcgt	3600
cgtgacccat ggcgatgcct gcttgccgaa	tatcatggtg gaaaatggcc gcttttctgg	3660
attcatcgac tgtggccggc tgggtgtggc	ggaccgctat caggacatag cgttggctac	3720
ccgtgatatt gctgaagagc ttggcggcga	atgggctgac cgcttcctcg tgctttacgg	3780
tatcgccgct cccgattcgc agcgcacgc	cttctatcgc cttcttgacg agttcttctg	3840
agcgggactc tggggttcga aatgaccgac	caagcgacgc ccaacctgcc atcacgagat	3900

ttcgattcca	ccgccgcctt	ctatgaaagg	ttgggcttcg	gaatcgtttt	ccgggacgcc	3960
ggctggatga	tcctccagcg	cggggatctc	atgctggagt	tcttcgcca	ccccaacttg	4020
tttattgcag	cttataatgg	ttacaaataa	agcaatagca	tcacaaattt	cacaaataaa	4080
gcattttttt	cactgcattc	tagttgtggt	ttgtccaaac	tcatcaatgt	atcttatcat	4140
gtctgtatac	cgtcgacctc	tagctagagc	ttggcgtaat	catggtcata	gctgtttcct	4200
gtgtgaaatt	gttatccgct	cacaattcca	cacaacatac	gagccggaag	cataaagtgt	4260
aaagcctggg	gtgcctaata	agtgaagctaa	ctcacattaa	ttgcgttgcg	ctcactgccc	4320
gctttccagt	cgggaaacct	gtcgtgccag	ctgcattaat	gaatcggcca	acgcgcgggg	4380
agaggcggtt	tgcgtatttg	gcgctcttcc	gcttcctcgc	tactgactc	gctgcgctcg	4440
gtcgttcggc	tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatacg	gttatccaca	4500
gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	ggccaggaac	4560
cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	4620
aaaaatcgac	gctcaagtca	gagggtggcg	aaccgcagag	gactataaag	ataccaggcg	4680
tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	4740
ctgtccgcct	ttctcccttc	gggaagcggt	gcgctttctc	aatgctcacg	ctgtagggtat	4800
ctcagttcgg	tgtaggctcg	tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	4860
cccgaccgct	gcgccttata	cggtaactat	cgtcttgagt	ccaacccggt	aagacacgac	4920
ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	4980
gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	agtatttggt	5040
atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	5100
aaacaaacca	ccgctggtag	cgggtggtttt	tttgtttgca	agcagcagat	tacgcgcaga	5160
aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaac	5220
gaaaactcac	gttaagggat	tttggtcatg	agattatcaa	aaaggatctt	cacctagatc	5280
cttttaaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttggctc	5340
gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	5400
tccatagtgt	cctgactccc	cgctcgttag	ataactacga	tacgggaggg	cttaccatct	5460
ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	tttatcagca	5520
ataaaccagc	cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	5580
atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	5640

cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggatatggct	5700
tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa	5760
aaagcggtta gtccttcggg tcctccgacg gttgtcagaa gtaagttggc cgcagtgtta	5820
tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc	5880
ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat ggggcgaccg	5940
agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa	6000
gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg	6060
agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc	6120
accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg	6180
gcgacacgga aatggtgaat actcatactc ttcctttttc aatattattg aagcatttat	6240
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata	6300
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tcgacggatc gggagatctc	6360
ccgatcccct atggctgact ctcagtacaa tctgctctga tgccgcatag ttaagccagt	6420
atctgctccc tgcttgtgtg ttggaggctg ctgagtagtg cgcgagcaaa atttaagcta	6480
caacaaggca aggcttgacc gacaattgca tgaagaatct gcttaggggtt aggcgttttg	6540
cgctgcttcg cgatgtacgg gccagatata cgcgttgaca ttgattattg actagttatt	6600
aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc cgcgttacat	6660
aacttacggg aaatggcccc cctggctgac cgcccaacga cccccgcca ttgacgtcaa	6720
taatgacgta tgttcccata gtaacgcca tagggacttt ccattgacgt caatgggtgg	6780
actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg ccaagtacgc	6840
cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag tacatgacct	6900
tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt accatggtga	6960
tgcggttttg gcagtacatc aatgggcgtg gatagcgggt tgactcacgg ggatttccaa	7020
gtctccacc cattgacgtc aatgggagtt tgttttggca ccaaaatcaa cgggactttc	7080
caaaatgtcg taacaactcc gcccattga cgcaaattgg cggtaggcgt gtacgggtggg	7140
aggtctatat aagcagagct ctctggctaa ctagagaacc cactgcttac tggcttatcg	7200
aaattaatac gactcactat agggagaccc aagcttggtg cc	7242

<210> 11

<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Peptide

<220>  
<221> MISC\_FEATURE  
<222> (3)..(3)  
<223> Xaa is any amino acid

<220>  
<221> MISC\_FEATURE  
<222> (4)..(4)  
<223> Xaa is any amino acid

<220>  
<221> MISC\_FEATURE  
<222> (6)..(6)  
<223> Xaa is any amino acid

<220>  
<221> MISC\_FEATURE  
<222> (8)..(8)  
<223> Xaa is any amino acid

<400> 11

Gln Gln Xaa Xaa Arg Xaa Leu Xaa Gly Tyr  
1 5 10

<210> 12  
<211> 16  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Peptide

<220>  
<221> MISC\_FEATURE  
<222> (3)..(3)  
<223> Xaa is any amino acid

<220>  
<221> MISC\_FEATURE  
<222> (5)..(5)  
<223> Xaa is any amino acid

<220>  
<221> MISC\_FEATURE  
<222> (6)..(6)  
<223> Xaa is any amino acid

<220>  
 <221> MISC\_FEATURE  
 <222> (8)..(8)  
 <223> Xaa is any amino acid

<220>  
 <221> MISC\_FEATURE  
 <222> (11)..(11)  
 <223> Xaa is any amino acid

<220>  
 <221> MISC\_FEATURE  
 <222> (12)..(12)  
 <223> Xaa is any amino acid

<220>  
 <221> MISC\_FEATURE  
 <222> (15)..(15)  
 <223> Xaa is any amino acid

<220>  
 <221> MISC\_FEATURE  
 <222> (16)..(16)  
 <223> Xaa is any amino acid

<220>  
 <221> MISC\_FEATURE  
 <222> (17)..(17)  
 <223> Xaa is any amino acid

<400> 12

Gln	Gln	Xaa	Glu	Xaa	Xaa	Leu	Xaa	Gly	Tyr	Xaa	Xaa	Trp	Ile	Xaa	Xaa
1				5					10					15	

<210> 13  
 <211> 1833  
 <212> DNA  
 <213> Synthetic Construct

<220>  
 <221> CDS  
 <222> (1)..(1833)

<400> 13																
atg	gat	gac	gac	tcc	tac	gtg	ccc	ggg	ttt	gag	gac	tcg	gag	gcg	ggt	48
Met	Asp	Asp	Asp	Ser	Tyr	Val	Pro	Gly	Phe	Glu	Asp	Ser	Glu	Ala	Gly	
1				5					10					15		
tca	gcc	gac	tcc	tac	acc	agc	cgc	ccc	tct	ctg	gac	tca	gac	gtt	tcc	96
Ser	Ala	Asp	Ser	Tyr	Thr	Ser	Arg	Pro	Ser	Leu	Asp	Ser	Asp	Val	Ser	
			20					25					30			
ctg	gag	gag	gac	cgg	gag	agt	gcc	cgg	cga	gaa	gtg	gag	agt	cag	gct	144

Leu	Glu	Glu	Asp	Arg	Glu	Ser	Ala	Arg	Arg	Glu	Val	Glu	Ser	Gln	Ala		
		35					40					45					
cag	cag	cag	ctg	gaa	aga	gcc	aag	cac	aaa	cct	gtg	gca	ttt	gct	gtg	192	
Gln	Gln	Gln	Leu	Glu	Arg	Ala	Lys	His	Lys	Pro	Val	Ala	Phe	Ala	Val		
	50					55					60						
agg	acc	aat	gtc	agc	tac	tgt	gga	gtt	ctg	gat	gag	gaa	tgc	cca	gtc	240	
Arg	Thr	Asn	Val	Ser	Tyr	Cys	Gly	Val	Leu	Asp	Glu	Glu	Cys	Pro	Val		
65					70					75					80		
cag	ggc	tct	gga	gtc	aac	ttc	gag	gcc	aaa	gat	ttt	ctg	cac	att	aaa	288	
Gln	Gly	Ser	Gly	Val	Asn	Phe	Glu	Ala	Lys	Asp	Phe	Leu	His	Ile	Lys		
				85					90					95			
gag	aag	tac	agc	aat	gac	tgg	tgg	atc	ggg	agg	cta	gtg	aaa	gaa	ggc	336	
Glu	Lys	Tyr	Ser	Asn	Asp	Trp	Trp	Ile	Gly	Arg	Leu	Val	Lys	Glu	Gly		
			100					105					110				
ggc	gat	att	gcc	ttc	atc	ccc	agc	ccc	caa	cgc	ctg	gag	agc	atc	cgg	384	
Gly	Asp	Ile	Ala	Phe	Ile	Pro	Ser	Pro	Gln	Arg	Leu	Glu	Ser	Ile	Arg		
		115				120						125					
ctc	aaa	cag	gaa	cag	aag	gcc	agg	aga	tcc	ggg	aac	cct	tcc	agc	ctg	432	
Leu	Lys	Gln	Glu	Gln	Lys	Ala	Arg	Arg	Ser	Gly	Asn	Pro	Ser	Ser	Leu		
	130					135					140						
agt	gac	att	ggc	aac	cga	cgt	tcc	cct	cct	cca	tct	cta	gcc	aag	cag	480	
Ser	Asp	Ile	Gly	Asn	Arg	Arg	Ser	Pro	Pro	Pro	Ser	Leu	Ala	Lys	Gln		
145					150					155					160		
aag	caa	aag	cag	gcg	gaa	cat	gtc	ccc	ccg	tat	gat	gtg	gtg	ccc	tcc	528	
Lys	Gln	Lys	Gln	Ala	Glu	His	Val	Pro	Pro	Tyr	Asp	Val	Val	Pro	Ser		
				165					170					175			
atg	cgg	cct	gtg	gtg	ctg	gtg	gga	ccc	tct	ctg	aaa	ggc	tat	gag	gtc	576	
Met	Arg	Pro	Val	Val	Leu	Val	Gly	Pro	Ser	Leu	Lys	Gly	Tyr	Glu	Val		
			180					185					190				
aca	gac	atg	atg	cag	aag	gct	ctc	ttc	gac	ttc	ctt	aaa	cac	agg	ttt	624	
Thr	Asp	Met	Met	Gln	Lys	Ala	Leu	Phe	Asp	Phe	Leu	Lys	His	Arg	Phe		
		195					200					205					
gat	ggc	agg	atc	tcc	atc	acc	cgc	gtc	acg	gct	gac	ctc	tca	ctg	gcc	672	
Asp	Gly	Arg	Ile	Ser	Ile	Thr	Arg	Val	Thr	Ala	Asp	Leu	Ser	Leu	Ala		
	210					215					220						
aag	cgc	tct	gtg	ctc	aac	aat	cct	ggc	aag	agg	acc	atc	atc	gag	cgc	720	
Lys	Arg	Ser	Val	Leu	Asn	Asn	Pro	Gly	Lys	Arg	Thr	Ile	Ile	Glu	Arg		
225					230					235				240			
tct	tct	gcc	cgc	tcc	agc	att	gct	gag	gtg	cag	agt	gag	att	gag	cgc	768	
Ser	Ser	Ala	Arg	Ser	Ser	Ile	Ala	Glu	Val	Gln	Ser	Glu	Ile	Glu	Arg		
				245					250					255			
ata	ttc	gag	ctg	gcc	aaa	tcc	ctg	cag	cta	gtg	gtg	ttg	gat	gct	gac	816	
Ile	Phe	Glu	Leu	Ala	Lys	Ser	Leu	Gln	Leu	Val	Val	Leu	Asp	Ala	Asp		

260										265										270										
acc	atc	aac	cac	cca	gca	cag	cta	gcc	aag	acc	tca	ctg	gcc	ccc	atc	864														
Thr	Ile	Asn	His	Pro	Ala	Gln	Leu	Ala	Lys	Thr	Ser	Leu	Ala	Pro	Ile															
275					280					285																				
atc	gtc	ttc	gtc	aaa	gtg	tcc	tcg	cca	aag	gta	ctg	cag	cga	ctg	atc	912														
Ile	Val	Phe	Val	Lys	Val	Ser	Ser	Pro	Lys	Val	Leu	Gln	Arg	Leu	Ile															
290					295					300																				
cgc	tcc	agg	ggg	aag	tcc	cag	atg	aag	cac	ctc	act	gta	cag	atg	atg	960														
Arg	Ser	Arg	Gly	Lys	Ser	Gln	Met	Lys	His	Leu	Thr	Val	Gln	Met	Met															
305	310					315					320																			
gca	tat	gat	aag	ctg	ggt	cag	tgc	cca	cct	gag	tca	ttt	gat	gtg	att	1008														
Ala	Tyr	Asp	Lys	Leu	Val	Gln	Cys	Pro	Pro	Glu	Ser	Phe	Asp	Val	Ile															
325					330					335																				
ctg	gat	gag	aac	cag	ctg	gac	gac	gcc	tgt	gag	cac	cta	gct	gaa	tac	1056														
Leu	Asp	Glu	Asn	Gln	Leu	Asp	Asp	Ala	Cys	Glu	His	Leu	Ala	Glu	Tyr															
340					345					350																				
cta	gag	ggt	tac	tgg	cgc	gct	acc	cac	cac	cca	gca	ccg	ggc	ccc	ggg	1104														
Leu	Glu	Val	Tyr	Trp	Arg	Ala	Thr	His	His	Pro	Ala	Pro	Gly	Pro	Gly															
355					360					365																				
atg	ctg	ggt	ccg	ccc	agt	gcc	atc	cct	gga	ctt	cag	aac	cag	cag	ctg	1152														
Met	Leu	Gly	Pro	Pro	Ser	Ala	Ile	Pro	Gly	Leu	Gln	Asn	Gln	Gln	Leu															
370					375					380																				
ctg	ggg	gag	cga	ggt	gag	gag	cat	tca	ccc	ctg	gag	cgg	gac	agt	ttg	1200														
Leu	Gly	Glu	Arg	Gly	Glu	Glu	His	Ser	Pro	Leu	Glu	Arg	Asp	Ser	Leu															
385	390					395					400																			
atg	ccc	tcg	gat	gag	gcc	agt	gag	agc	tcc	cgc	cag	gcc	tgg	acc	gga	1248														
Met	Pro	Ser	Asp	Glu	Ala	Ser	Glu	Ser	Ser	Arg	Gln	Ala	Trp	Thr	Gly															
405					410					415																				
tct	tca	cag	cgc	agc	tcc	cgc	cat	ctg	gag	gag	gac	tat	gca	gat	gcc	1296														
Ser	Ser	Gln	Arg	Ser	Ser	Arg	His	Leu	Glu	Glu	Asp	Tyr	Ala	Asp	Ala															
420					425					430																				
tac	cag	gac	ctg	tac	cag	cct	cac	cgt	caa	cac	acc	tcg	ggg	cta	ccc	1344														
Tyr	Gln	Asp	Leu	Tyr	Gln	Pro	His	Arg	Gln	His	Thr	Ser	Gly	Leu	Pro															
435					440					445																				
agt	gct	aac	ggg	cat	gac	ccc	caa	gac	cgg	ctc	cta	gcc	cag	gac	tcg	1392														
Ser	Ala	Asn	Gly	His	Asp	Pro	Gln	Asp	Arg	Leu	Leu	Ala	Gln	Asp	Ser															
450					455					460																				
gag	cat	gac	cac	aat	gac	cgg	aac	tgg	cag	cgt	aac	cgg	cct	tgg	cct	1440														
Glu	His	Asp	His	Asn	Asp	Arg	Asn	Trp	Gln	Arg	Asn	Arg	Pro	Trp	Pro															
465					470					475					480															
aag	gac	agc	tac	gaa	ttc	gcc	aaa	gaa	agg	gag	cgg	gtg	gag	aac	cgg	1488														
Lys	Asp	Ser	Tyr	Glu	Phe	Ala	Lys	Glu	Arg	Glu	Arg	Val	Glu	Asn	Arg															
485					490					495																				

cg	gc	tt	ct	aa	ct	cg	cg	ca	ca	ca	at	ga	cg	ga	ct	1536
Arg	Ala	Phe	Leu	Lys	Leu	Arg	Arg	Gln	Gln	Gln	Ile	Glu	Arg	Glu	Leu	
			500					505					510			
aa	gg	ta	at	ga	tg	at	tc	aa	gc	ga	ga	ga	gt	at	ct	1584
Asn	Gly	Tyr	Met	Glu	Trp	Ile	Ser	Lys	Ala	Glu	Glu	Val	Ile	Leu	Ala	
		515					520					525				
ga	ga	ga	ac	ga	gt	ga	ca	ag	ca	ca	cc	tt	ga	gg	gc	1632
Glu	Asp	Glu	Thr	Asp	Val	Glu	Gln	Arg	His	Pro	Phe	Asp	Gly	Ala	Leu	
	530					535					540					
cg	ag	gc	ac	at	aa	aa	ag	ac	ga	ct	ct	ca	cc	ga		1680
Arg	Arg	Ala	Thr	Ile	Lys	Lys	Ser	Lys	Thr	Asp	Leu	Leu	His	Pro	Glu	
545					550				555						560	
ga	gc	ga	ga	ca	ct	gc	ga	at	gc	tc	gt	gg	tc	cc	tt	1728
Glu	Ala	Glu	Asp	Gln	Leu	Ala	Asp	Ile	Ala	Ser	Val	Gly	Ser	Pro	Phe	
				565				570						575		
gc	ca	gc	ag	at	aa	ag	gc	aa	ct	ga	aa	tc	ag	tt	tt	1776
Ala	Arg	Ala	Ser	Ile	Lys	Ser	Ala	Lys	Leu	Glu	Asn	Ser	Ser	Phe	Phe	
			580					585					590			
ca	aa	aa	ga	ag	ag	at	cg	tt	ta	at	cg	cg	at	gt	aa	1824
His	Lys	Lys	Glu	Arg	Arg	Met	Arg	Phe	Tyr	Ile	Arg	Arg	Met	Val	Lys	
		595					600					605				
ac	ca	ta														1833
Thr	Gln															
		610														
<210>	14															
<211>	610															
<212>	PRT															
<213>	Synthetic Construct															
<400>	14															
Met	Asp	Asp	Asp	Ser	Tyr	Val	Pro	Gly	Phe	Glu	Asp	Ser	Glu	Ala	Gly	
1				5					10					15		
Ser	Ala	Asp	Ser	Tyr	Thr	Ser	Arg	Pro	Ser	Leu	Asp	Ser	Asp	Val	Ser	
			20					25					30			
Leu	Glu	Glu	Asp	Arg	Glu	Ser	Ala	Arg	Arg	Glu	Val	Glu	Ser	Gln	Ala	
		35					40					45				
Gln	Gln	Gln	Leu	Glu	Arg	Ala	Lys	His	Lys	Pro	Val	Ala	Phe	Ala	Val	
		50				55					60					



Arg Thr Asn Val Ser Tyr Cys Gly Val Leu Asp Glu Glu Cys Pro Val  
 65 70 75 80

Gln Gly Ser Gly Val Asn Phe Glu Ala Lys Asp Phe Leu His Ile Lys  
 85 90 95

Glu Lys Tyr Ser Asn Asp Trp Trp Ile Gly Arg Leu Val Lys Glu Gly  
 100 105 110

Gly Asp Ile Ala Phe Ile Pro Ser Pro Gln Arg Leu Glu Ser Ile Arg  
 115 120 125

Leu Lys Gln Glu Gln Lys Ala Arg Arg Ser Gly Asn Pro Ser Ser Leu  
 130 135 140

Ser Asp Ile Gly Asn Arg Arg Ser Pro Pro Pro Ser Leu Ala Lys Gln  
 145 150 155 160

Lys Gln Lys Gln Ala Glu His Val Pro Pro Tyr Asp Val Val Pro Ser  
 165 170 175

Met Arg Pro Val Val Leu Val Gly Pro Ser Leu Lys Gly Tyr Glu Val  
 180 185 190

Thr Asp Met Met Gln Lys Ala Leu Phe Asp Phe Leu Lys His Arg Phe  
 195 200 205

Asp Gly Arg Ile Ser Ile Thr Arg Val Thr Ala Asp Leu Ser Leu Ala  
 210 215 220

Lys Arg Ser Val Leu Asn Asn Pro Gly Lys Arg Thr Ile Ile Glu Arg  
 225 230 235 240

Ser Ser Ala Arg Ser Ser Ile Ala Glu Val Gln Ser Glu Ile Glu Arg  
 245 250 255

Ile Phe Glu Leu Ala Lys Ser Leu Gln Leu Val Val Leu Asp Ala Asp  
 260 265 270

Thr Ile Asn His Pro Ala Gln Leu Ala Lys Thr Ser Leu Ala Pro Ile  
 275 280 285

Ile Val Phe Val Lys Val Ser Ser Pro Lys Val Leu Gln Arg Leu Ile

290		295		300
Arg Ser Arg Gly Lys Ser Gln Met Lys His Leu Thr Val Gln Met Met				
305		310		315 320
Ala Tyr Asp Lys Leu Val Gln Cys Pro Pro Glu Ser Phe Asp Val Ile				
	325		330	335
Leu Asp Glu Asn Gln Leu Asp Asp Ala Cys Glu His Leu Ala Glu Tyr				
	340		345	350
Leu Glu Val Tyr Trp Arg Ala Thr His His Pro Ala Pro Gly Pro Gly				
	355		360	365
Met Leu Gly Pro Pro Ser Ala Ile Pro Gly Leu Gln Asn Gln Gln Leu				
	370		375	380
Leu Gly Glu Arg Gly Glu Glu His Ser Pro Leu Glu Arg Asp Ser Leu				
385		390		395 400
Met Pro Ser Asp Glu Ala Ser Glu Ser Ser Arg Gln Ala Trp Thr Gly				
	405		410	415
Ser Ser Gln Arg Ser Ser Arg His Leu Glu Glu Asp Tyr Ala Asp Ala				
	420		425	430
Tyr Gln Asp Leu Tyr Gln Pro His Arg Gln His Thr Ser Gly Leu Pro				
	435		440	445
Ser Ala Asn Gly His Asp Pro Gln Asp Arg Leu Leu Ala Gln Asp Ser				
	450		455	460
Glu His Asp His Asn Asp Arg Asn Trp Gln Arg Asn Arg Pro Trp Pro				
465		470		475 480
Lys Asp Ser Tyr Glu Phe Ala Lys Glu Arg Glu Arg Val Glu Asn Arg				
	485		490	495
Arg Ala Phe Leu Lys Leu Arg Arg Gln Gln Gln Ile Glu Arg Glu Leu				
	500		505	510
Asn Gly Tyr Met Glu Trp Ile Ser Lys Ala Glu Glu Val Ile Leu Ala				
	515		520	525

Glu Asp Glu Thr Asp Val Glu Gln Arg His Pro Phe Asp Gly Ala Leu  
530 535 540

Arg Arg Ala Thr Ile Lys Lys Ser Lys Thr Asp Leu Leu His Pro Glu  
545 550 555 560

Glu Ala Glu Asp Gln Leu Ala Asp Ile Ala Ser Val Gly Ser Pro Phe  
565 570 575

Ala Arg Ala Ser Ile Lys Ser Ala Lys Leu Glu Asn Ser Ser Phe Phe  
580 585 590

His Lys Lys Glu Arg Arg Met Arg Phe Tyr Ile Arg Arg Met Val Lys  
595 600 605

Thr Gln  
610